

Dynamic Grooming in IP over WDM Networks: A Study with Realistic Traffic based on GANCLES Simulation Package

E. Salvadori, R. Lo Cigno

Dipartimento di Informatica e Telecomunicazioni
Università di Trento
Via Sommarive, 14, 38050 Povo, Trento, Italy
e-mail: {salvador, locigno}@dit.unitn.it

Z. Zsóka

Department of Telecommunications
Budapest University of Technology and Economics
Magyar tudósok körútja 2, 1117 Budapest, Hungary
e-mail: zsoka@hit.bme.hu

Abstract—Dynamic grooming capabilities lies at the hearth of many envisaged scenarios for IP over Optical networks, but studies on its performance are still in their infancy. This work addresses two fundamental aspects of the problem.

First of all it presents a novel tool for the study of IP over Optical networks. The tool, freely available on-line, is a network level simulator named GANCLES that includes several innovative features allowing the study of realistic scenarios in IP over Optical networking, making it an ideal tool for Traffic Engineering purposes. GANCLES architecture enables the simulation of dynamic traffic grooming on top of a realistic network model that correctly describes the logical interaction between the optical and the IP layer, i.e., the mutual relationship between routing algorithms and lightpath assignment procedures at the optical layer and routing at the IP layer. Adding or removing lightpaths changes the logical IP topology, which affects IP routing and traffic patterns. The simulator allows for the description of Overlay, Augmented or Peer IP over Optical architectures, depending on the amount of information shared between the IP and optical domain.

Second it analyzes and discusses several performance indices and aspects of different grooming policies in the IPO Overlay model, using different traffic models, some of them including elasticity of best effort traffic. Both regular and mesh topologies are analyzed, and results clearly show that the correct evaluation of dynamic grooming policies in IPO networks requires a sophisticated level of modeling, since simplistic assumptions like Poisson traffic, or the incorrect representation of the interaction of IP and Optical control planes, may induce misleading results.

I. INTRODUCTION AND RELATED WORK

IP over WDM networks[1] are spreading very fast and even tier 1 providers have begun to encapsulate IP packets directly in the optical layer, avoiding the use of sophisticated middle layers such as ATM (Asynchronous Trans-

This work is supported by the Italian Ministry of Education and Research (MIUR) through the GRID.IT and ADONIS projects, and by the Hungarian Italian Intergovernmental S&T Cooperation Programme '04-'07 (I-17/03); GANCLES is developed under the E-NEXT NoE umbrella.

fer Mode) or SONET/SDH (Synchronous Optical Network/Synchronous Digital Hierarchy).

The optical layer is managed through protocols like GMPLS (Generalized Multi-Protocol Label Switching)[2] or ASON (Automatically Switched Optical Network)[3], while the IP layer is either integrated in a MPLS framework or uses standard intra-AS IP routing protocols such as OSPF or IS-IS.

Dynamic grooming of IP traffic over a wavelength routed optical network means that the two routing layers (IP and optical) interact, with deep impacts on Traffic Engineering (TE) and QoS provisioning. The interaction nature depends on the grooming algorithm, as well as on the amount of information (if any) exchanged between the two layers. This situation is very complex and its study is normally done via simulations with a modeling effort to reduce the problem complexity, e.g., without simulating packet level traffic, but with fluid models. Sivalingam et al. have presented an *ns-2* based simulation tool for performance studies of WDM networks[4]. This simulator does not consider the problem of grooming and the WDM management layer is seen as a logical layer on top of an IP network (the standard *ns-2* network layer) building virtual circuits on the packed switched routing layer, thus its “philosophy” is completely different from GANCLES.

From the perspective of performance analysis, a few recent studies started considering the problem of dynamic grooming[5], [6], [7], [8], [9], but, to the best of our knowledge, only simplistic models of the network dynamics associated to CBR Poisson traffic sources were used in these studies. More traditional studies such as [5], [10], [11], [12] are based on the simplifying assumption that the traffic is static and “circuit-like.” In a previous contribution[13] we discussed in detail the impact of realistic traffic models on dynamic grooming, showing the inherent interaction between the IP and the optical layer and its effect on the overall performance of the network. Traffic flows in this models

share the resources on a virtual topology path following the max-min fairness criterion [14], thus mimicking the ideal behavior of a bundle of TCP connections.

The contribution of this paper is twofold. In the first part we present GANCLES, making it available to the community for use in research and applications. A major achievement in GANCLES is that it allows the study of the interaction between four major drivers of the overall network performance:

- The optical layer Routing and Wavelength Assignment (RWA) algorithms;
- The IP level routing;
- The grooming policies used to mediate between the IP level and the optical level;
- The adaptability of current data applications, that, being based on TCP, modify the transmission rate following the availability of resources.

In particular the clear separation of the IP and Optical control planes, enables the correct definition of the IPO model (peer, augmented, overlay) [15] addressed with the simulation experiment, and the presence of elastic traffic makes the experiments representative of IP over Optical, relaxing the usual approximation based on CBR, bandwidth guaranteed traffic.

In the second part we present results for different, albeit not entirely new, dynamic grooming policies, comparing their behavior and highlighting the scenarios, conditions and performance indices that require proper, in-depth studies for the assessment of dynamic grooming policies and architectures. In this work we consider only the Overlay model, in part because it is impossible to include all different facets in a single work and in part because we think the Overlay model is the only one which is technologically foreseeable in the next future, and for this reason, the most interesting from the implementation point of view.

The remaining part of the paper is organized as follows. Sect. II describes thoroughly the simulation environment and the innovative features we introduced to study dynamic IP over WDM networks. In Sect. III we define some performance figures of IPO networks which can be studied by using GANCLES. Sect. IV is devoted to the discussion of results, both in a ring and on the NSF topology; Sect. V ends the paper and discusses future work enabled by GANCLES.

II. THE TOOL FEATURES AND ARCHITECTURE

GANCLES [16] is an event-driven asynchronous simulator derived from ANCLES [17]. ANCLES has gradually evolved over the years to allow the simulation of elastic connections over IP networks, with a flow-level granularity, as described in [18]. A separate extension allows the study of different lightpath-level granularity in ASON-based wavelength-routed networks [19].

GANCLES integrates the two network layers (from now on: the *data-layer* and the *optical-layer*), thus allowing the

in-depth study of the interaction between them when a multi-layer network environment such as IP over WDM is considered. The objective of the simulator is to give researchers a useful tool to study new algorithms and protocols, to analyze network performance, to implement traffic engineering criteria, and to design QoS provisioning means in this multi-layer environment.

The simulator includes advanced tools to perform statistical analysis based on the “batch-means” technique [20]. Simulation experiments are stopped when the desired accuracy is reached on a selected set of performance parameters.

The tool allows the computation of a large number of performance indexes, both for the entire network and for selected traffic relations. Some of them are illustrated in more detail in Sect. III.

The network models simulated by the tool are composed of instances of three basic entities.

- **NODES**, which perform the routing functions at the IP and at the WDM layer, and implement the CAC and the grooming algorithms; NODES can be either pure OXCs (Optical Crossconnect), switching entire lightpaths, or include an IP router on top of the OXC, allowing for data-layer traffic injections/extraction and performing grooming operations; these nodes are named G-OXC (Grooming-OXC).
- **CHANNELS**, that accommodate the information transfer between either adjacent NODES or USER-NODE pairs; a CHANNEL can accommodate up to W independent lightpaths.
- **USERS**, that acts as sources and sinks for the traffic flowing through the network; USERS can be both at the optical-layer, generating “circuit-like” requests of entire lightpaths, and connected directly to OXCs, or at the data-layer, connected to G-OXCs only and generating sub-wavelength requests that can follow one of three different “models:” i) traditional circuit-like requests, ii) Time Based (TB) best-effort requests, and iii) Data Based (DB) best-effort requests (See Sect II-C for further details).

Simulations are specified with a formal grammar inherited from [17] and named ND. The number of CHANNELS and their data rates are expressed in number of fiber per link, number of wavelengths per fiber and finally, transmission capacity in Gbit/s per wavelength. The NODE architecture is described in detail discussing the interaction between the data-layer and optical-layer in Sect. II-B

A. IP over Optical architectures

One of the most important feature of GANCLES is enabling the implementation of dynamic grooming algorithms which refer to different IP over Optical (IPO) architectures as defined by RFC 3717 [15]. According to this RFC, when considering the *control plane* of an IPO network,

three different interconnection models can be envisioned according to the amount of information exchanged between the IP and optical layer: *Overlay, Augmented and Peer*.

In the *peer* model, IP routers and OXC are considered peer network elements, thus the topology and other network information are completely shared by a unified control plane. In the *overlay* model, each layer performs its own routing functions since no information is exchanged between them. An intermediate architecture between these two is the *augmented* model, where some aggregated information from one routing instance is passed to the other, in general only from the optical to the IP layer, in order to allow this latter to take more informed decisions when submitting requests for additional resources.

The peer and the augmented model are appealing because sharing the knowledge base between the two layers allows running an integrated routing function, using, for instance, an auxiliary graph, as done in [5]. The integrated management enables a better usage of the overall network resources. However, both models seem not feasible in the near term due to the tight integration between the two levels and scalability issues regarding the amount of exchanged information. The overlay model is instead technically feasible, since it only requires the definition of a clear interface between the IP and optical level and dynamic lightpath capabilities in the optical level, which are being experimented in laboratories and research projects [21].

Due to its specific implementation characteristics, GANCLES enables the study of scenarios with full, partial or no exchange of information between the data-layer and the optical-layer, therefore allowing the description of any of the grooming algorithms proposed so far in literature.

B. Physical and Logical Topology Management and Interaction

As mentioned before, in GANCLES nodes can be pure OXCs, that switch entire lightpaths from an ingress port to an egress port, or they can be G-OXCs that support sub-wavelength traffic flows and multiplex them onto wavelength channels through a grooming fabric. OCXs and G-OXCs can be mixed freely into a simulation experiment. It is possible to have both *full opaque* or *full transparent* crossconnects. Opaque OXCs allows full wavelength conversion; transparent OXCs have no conversion capabilities. Partially opaque OXCs, with limited conversion capabilities, are being implemented. A G-OXC is *also* a router, hence the transit traffic (not terminated in the router), can be groomed with incoming traffic. Sub-wavelength traffic can be generated and received only in G-OXCs.

When considering a multi-layer environment, with connections flowing between IP level nodes through an optical network, we need to distinguish the *physical* topology and the *logical* topology. The latter one is made of all the lightpaths established between G-OXCs over the physical

topology according to some optical-level routing algorithms. The logical topology is used for routing at the data-layer (IP) and it is modified each time the grooming management entity triggers the establishment or release of some lightpaths.

Simulation in GANCLES are driven by the USERS, which collect requests from their associated call generators and forward them to the network, while acting as destinations for the connections coming from remote users.

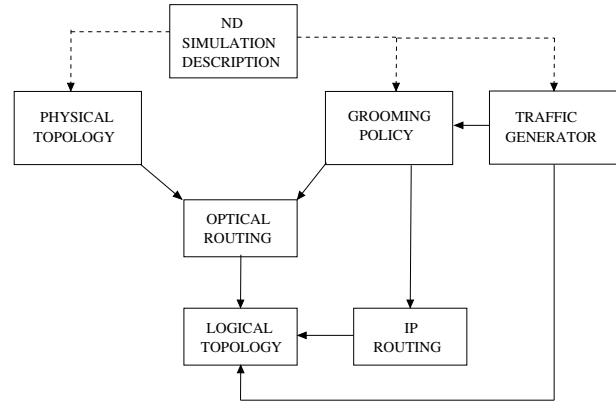


Fig. 1. Logical interaction between different high-level modules in GANCLES, the management of the optical-layer is mediated by the grooming strategies and algorithms

Fig. 1 represents the interaction between different GANCLES parts. A simulation starts after GANCLES has acquired the simulation experiment description in ND. The description includes: (i) the network topology in terms of a weighted graph connecting USERS and NODES through CHANNELS; (ii) the traffic relations between USERS and the statistical characterization of sources; (iii) the selection of the routing, CAC and grooming algorithms adopted for the simulation run; (iv) a number of options concerning both the network operation and the simulation session management; (v) the performance indices to be measured.

When the simulation starts an optical-layer function independent from the selected RWA algorithm defines the set of available physical paths for each pair of OXCs/G-OXCs ordering them according to some specific criterion. Only this set can be used by the routing algorithm, reducing the routing problem complexity.

Every time a new lightpath is added or removed from the network by the optical-layer, the data-layer (logical) topology is changed. As done at the optical-layer, also in this case the set of possible logical paths between each G-OXC (router) pair is computed following a user-specified criterion. This task is extremely critical, because path-computing is very time consuming, hence the path selection criterion must be carefully defined depending on the data-layer routing algorithm selected (e.g., only one path need to be computed if Fixed Shortest Path routing is used).

Notice that the overall setup (and performance) of IPO networks is heavily influenced by technology constraints. We already mentioned opaque or transparent OXCs, but other constraints, such as simplex- or duplex-only lightpath management also come into play at the optical-layer. Similar constraints may arise if TE techniques are used at the data-layer. GANCLES allows the simulation of an arbitrary mix of these constraints.

Each time a USER generates a connection request, the grooming entity decides whether: (i) route it over the current topology; or (ii) ask the optical-layer to open one or more lightpaths (thus modifying the logical topology) and then route the request at the data-layer. In the first case, a data-layer CAC algorithm (if any) is executed for each of the paths being considered by the data-layer routing algorithm; if no path is found to route the incoming request in agreement with its QoS requirements, the request is dropped. In the second case, after the logical topology modification, the same data-layer routing and CAC are applied. As a general rule, in this case a connection can be refused only if the optical-layer was not able to modify the logical topology meeting the grooming algorithm requirements. For each request the grooming procedure may require an arbitrary mix of actions of type (i) and (ii), depending on the complexity of the algorithm implemented and on the model (peer, augmented, overlay) assumed for the IPO.

The interaction between the optical-layer and the data-layer lies at the core of dynamic grooming problems. This interaction is described in GANCLES (as in real networks) by the optical/IP control interface within G-OXCs. This simple and clear, though realistic, interface implementation is one of the innovative features of GANCLES.

Modifying the logical topology, the question arises whether all the traffic is re-routed or if only new connections follow the new available routes. GANCLES presently assumes the second option, but its modification for re-routing is trivial.

When an active connection terminates, the corresponding resources in the logical topology are released. This operation can lead to the release of some lightpaths if they are not carrying traffic anymore. In this case, paths at the IP level need to be recalculated again over the new logical topology. The specific lightpath releasing criteria (e.g., a given time-lapse without traffic) can be specified by GANCLES user.

GANCLES implicitly assumes the utilization of a separated control plane (e.g., GMPLS) to keep each node informed of the network status. The control traffic is not considered in the performance measures.

C. Traffic Sources

The present release of GANCLES provides different types of call generators at both the data-layer and optical-layer. Each connection request is associated with the identifier of

the destination USER, which is chosen accordingly to the traffic relations specified for the experiment.

The main traffic models implemented for the data-layer USERS are: CBR, ON-OFF CBR, Uniform VBR, Video VBR or Best-Effort with TB and DB model (see [17]), while for the optical-layer USERS it is possible to generate lightpath requests according to Poisson traffic generators, but it is also possible to generate permanent and semi-permanent lightpaths (see [19]), a feature that enables adding constraints or static portions to the logical topology.

Since GANCLES has been developed mainly to study the interaction between IP (as data-layer) and the optical-layer, a more detailed explanation of the *elastic* traffic features of Best-Effort USERS is needed. As a general rule IP does not implement CAC functions and the congestion control is done reactively by TCP. As we discussed in [13] the elastic nature of present-day data applications cannot be disregarded if dynamic grooming is considered, because the feedback introduced by the closed-loop nature of TCP (and traffic aggregation does not destroy the feedback) has an enormous impact on the overall performance. We introduce two different models of elastic traffic. Both share the characteristic that a flow i arrives to the network with a backlog of data D_i to transmit. Both include some form of elasticity, though very different one another.

The first model, named *time-based* (TB), assumes that the elasticity is taken into account only reducing the transfer rate when congestion arises. The flow duration is determined when the flow arrives to the network, based on its backlog D_i and its peak transmission rate B_{M_i} (e.g., the access link speed) $\tau_i = \frac{D_i}{B_{M_i}}$. The effect of congestion is just that the throughput of flows is reduced, but their closing time is not affected. A consequence of this behavior is that the data actually transferred by a flow i is generally less than the “requested” amount D_i , thus reducing the actual network load and relieving congestion. This model is very simple and does not grab all the complexity of the closed-loop interaction between the sources and the network.

A more accurate model, named *data-based* (DB), assumes instead an ideal max-min sharing of the resources within the network at any given instant. Flows still arrive to the network with a backlog D_i and with a peak transmission rate B_{M_i} . The acceptance of a new flow will affect not only all the other flows on the same path, but indeed all the flows in the network, since the max-min fair share is completely recomputed updating the estimated closing time of all the flows in the network. The same applies when flows close, freeing network resources. This model includes the most important feature of elastic traffic, which is the positive feedback on the flows duration. The more congested is the network, the longer the accepted flows remain in the network.

Without a CAC, at high loads the network can become

instable, as the number of flows within the network can grow to infinity and their individual throughput goes to zero. To build a more realistic scenario, a second attribute has been introduced to characterize any IP flow i : a minimum requested rate b_{m_i} . If at any time the bandwidth assigned to flow i falls below it, then flow i closes and is counted as a “starved” flow, because the network was not able to guarantee its correct completion. The attributes b_{m_i} and B_{M_i} are included in some SLA (Service Level Agreement) at the IP/Optical interface (see [22] for initial works on Optical-SLA).

A new performance measure can therefore be introduced, called the *starvation probability* p_s , which complements more traditional metrics such as throughput, blocking probability, optical-layer overhead in opening and closing lightpaths, etc.

D. Routing Algorithms

Inheriting the terminology in [17], whatever criterion is used to order the paths at both the optical-layer or data-layer, a *primary path* is always defined for each pair of NODES. All the other allowed paths are referred to as *secondary paths*.

Data-layer Routing Several alternatives are available to route calls at data-level so as to take into account the dynamic load of the network. An important property of GANCLES is the possibility to simulate both source-based (e.g., MPLS-like) or hop-by-hop routing.

The following list enumerates the main routing algorithms implemented in GANCLES. The reader is referred to [17] for more details on this part and the relative references to the literature.

- *Single Path Routing*: only the primary path between the nodes is considered to route the connection. This is also known as Fixed Shortest Path.
- *Controlled & Uncontrolled Alternate Routing*: if there is no space for the connection along the primary path, the secondary ones are investigated with different constraints (see [23] for details).
- *Minimum Distance Routing*: for each source-destination pair, the path π is chosen that minimizes the following quantity: $C_\pi = \sum_{l \in \pi} \frac{1}{b_l}$ where b_l is the max-min fair bandwidth that is available to a new connection over link l belonging to path π (see [24] for details).
- *Widest-Shortest Routing*: this algorithm first identifies the minimum-hop-count paths and breaks ties by choosing, among the paths with the minimum hop count, the one with the maximum available bandwidth (see [25] for details).

Optical-layer Routing There are different static and dynamic algorithms implemented for the routing of optical-

layer requests. In the following only single-path routing algorithms are described, but the tool provides also for protective routing, with both *dedicated* or *shared* protection mechanisms (see [19] for more details).

If we use opaque OXCs, the wavelength continuity is necessary only in the transparent sections, i.e. the part of route connecting two opaque nodes. Allowing wavelength conversion in the opaque nodes implies that the wavelength assigned to the connection can be different in each transparent section. The main optical routing algorithms implemented in GANCLES are:

- *Fixed Shortest Path*: it routes the lightpath request always on the dedicated primary path between the endpoints.
- *Shortest-Widest Path*: it selects the paths with the largest number of available optical channels; if there are more possibilities, it routes the lightpath on the shortest among them.
- *Alternate Shortest Path*: it selects the shortest from those paths where there is at least one wavelength available.

Furthermore, Wavelength Assignment (WA) algorithms can be freely associated to any routing. WA algorithms include *Random* and *First-Fit*, and others can be added easily if required (details and references on optical-layer routing and RWA can be found in [19]).

E. Grooming Algorithms

Grooming policies are the ensemble of algorithms and protocols that takes the decisions regarding possible changes of the current logical topology each time a data-layer request arrives or leaves the network. When new logical links need to be installed, two factors must be determined: how many of them must be set up and between which OXCs.

The decision to route the incoming requests over the existing logical topology or to establish new lightpaths to create more room for them can lead to different network performances. A general analysis of different “grooming policies” is carried out in [5] under the hypothesis of bandwidth-guaranteed traffic within a peer IPO model. When *elastic traffic* is considered, there is no obvious upper limit to the possible number of flows which is routed onto the existing logical layer. In this case, the need for the establishment of new lightpaths must be introduced based on some suitable parameter. We introduce this parameter, called *optical opening threshold* τ_o , as a threshold on the throughput that the incoming request would achieve on the selected route, defined as a fraction of the peak rate B_{M_i} required by each flow i present on the path analyzed.

Three grooming solutions are implemented in the current version of the simulator:

- *Virtual-topology First (VirtFirst)*. This grooming algorithm aims to open new optical lightpaths only if the

current virtual (logical) topology does not have enough resources to carry the incoming request. In case of new lightpath setup, if it is successful, the IP request is routed over it, otherwise the request is routed based on the current logical topology.

- *Optical-level First (OptFirst)*. Each time a new data-level request arrives in some router, the G-OXC always attempts to set up a new lightpath in the optical layer, in order to route the request over it. If no free wavelengths are available, the IP router routes the incoming request over the current logical topology. Indeed, a logical topology is defined only when optical resources for the considered source-destination pair are exhausted.
- *HopConstrained (HopCons)*. In this case, the decision on opening a new connection is taken based on path-based constraints. Given a number of IP-level hops K , a new lightpath between source s and destination d is set up if it is impossible to find a path on the current logical topology with less than $K + 1$ hops and more than $B_{Mi} \cdot \tau_o$ bandwidth available. *HopCons* can be considered an intermediate policy between the previous two, in fact it behaves as *VirtFirst* for $K \rightarrow \infty$ and as *OptFirst* for $K = 0$. *HopCons* grooming is described in detail and its implications are discussed in [26].

Note that in all these cases if the source-destination pair is disconnected on the logical topology and no new lightpath can be installed between them, the incoming connection request must be refused even if it is elastic. This phenomenon is particularly evident at low/medium loads with grooming algorithms using aggressively the optical resources, as it was shown in [13]. Such unacceptable situations could be avoided by including a pre-defined spanning tree or any other basic logical topology to ensure the connectivity in the data-layer, a feature GANCLES is provided with, defining different pre-defined logical topologies (see [16]).

When a lightpath needs to be released because it is not carrying traffic, the simulator gives the possibility of delaying the closure using a timeout, called *optical closing timeout* τ_{cl} . When no traffic is carried over some lightpath, it is kept open for the timeout period and gets closed only if its state does not change. This parameter can be specified by the user and can be adapted to the traffic characteristics, e.g. it can be set according to the mean interarrival time of data-layer requests. This dependency on the load structure can be very useful to avoid excessive oscillations in the logical topology, which are notoriously harmful to IP routing. Using management information it can be easily implemented in real networks.

III. PERFORMANCE INDICES

As discussed in Sect. IV, the phenomena involved in routing/grooming elastic traffic are rather complex, and often far from intuitive. The following performance indices blend both user-perceived performance and network operation

costs, thus helping in understanding the global performance of the network.

p_b — Blocking probability. It is the probability that a flow is not accepted, either because of some CAC function decision or (in best-effort traffic) because no connectivity can be found between the source and the destination.

p_s — Starvation probability. It is the probability that a best-effort flow closes during its life because it is not receiving service with acceptable quality. A flow i closes and drops the network if its instantaneous throughput falls below b_{mi} .

T — The average throughput per flow.

$$T = \frac{1}{N_c} \sum_{i=1}^{N_c} T_i$$

where N_c is the number of observed flows (e.g., during a simulation). Notice that in a resource sharing environment this is not the average resource occupation divided by the number of flows, since flows have all the same weight in the average, regardless of their dimension.

R_c — Routing table change rate. Each time a new lightpath is established, some of the routing tables must be recomputed due to the new virtual topology. The rate of such changes is a good measure of the joint grooming and routing cost within the network.

N_h — Average number of IP hops per flow.

N_{lo} — Time weighted average number of links per optical path. Each lightpath is weighted by its holding time, so that lightpaths lasting longer are correctly accounted for.

U_d — Distance unfairness index.

$$U_d = \frac{\max_{0 < r < N} T^r - \min_{0 < r < N} T^r}{T}$$

measures if the resource assignment is fair with respect to physical distance. T^r is the average throughput calculated for node-pairs with hop distance r in the physical topology, while N is the number of nodes in the network. It ranges from zero to ∞ ; any value larger than one indicates unacceptable unfairness. This parameter is evaluated only for regular topologies and uniform traffic, since in other cases it is can be influenced by factors external to the grooming policy.

Many other unfairness indices can be defined, based on different flow characteristics, e.g., the flow granularity (see for instance [27]), or in irregular topologies based on the source-destination $s - d$ traffic relation. Distance based unfairness is however typical of elastic traffic and it can be interesting, as discussed in Sect. IV-A on a ring topology, to investigate whether grooming policies can relieve, at least in part, the skewed behavior induced by the max-min sharing criterion.

TABLE I
PARAMETERS USED THROUGHOUT THE SIMULATIONS

Simulation parameters		Value
Fibers per link		1
OXC's		transparent
Number of wavelengths per fiber	W	8 for 8-nodes Ring 4 for NSFNet
Data rate per wavelength	g	20 Gbit/s
Duplex lightpaths		
Request arrival process	λ	Variable with load
Request backlog distribution: exponential	η	12.5 Gbit/s
Requested peak rate	B_M	10 Gbit/s
Requested minimum rate	b_m	1 Gbit/s
optical opening threshold	τ_o	0.1
optical closing timeout	τ_{cl}	0
Traffic pattern		Uniform among G-OXC's
Data flows		mono-directional
IP routing		Fixed Shortest Path
optical-layer routing		Shortest-Widest for 8-Ring alternate shortest path for NSFNet
Maximum number of hops in HopCons	K	1
Confidence interval		1% of point estimate
Confidence level		99%

In general the goal of a grooming algorithm is maximizing T while minimizing p_s and U_d , while it is not always straightforward to define “a goal” for other performance parameters, since they can be subject to contrasting needs.

IV. A SAMPLE STUDY

In this work, we are interested in understanding the fundamental aspects of dynamic grooming in overlay IPO networks, and the inherent interaction with data-based, elastic, best-effort traffic. At the same time we highlight the features we have introduced in GANCLES.

As already mentioned we limit the study to overlay networks because of their inherent simplicity and because we deem they will be the first ones to be deployed. In overlay networks the control planes of the IP and optical network are completely separated and there is no information exchange between IP and optical routing entities: the two layers interact only through the lightpath setup or tear-down. Lightpaths setup is requested by the IP layer whenever it needs additional resources, and lightpaths are teared down when they are not needed anymore¹.

We investigate two different scenarios, that are complementary one another.

8-Ring — This is an 8-node (all G-OXC's) bidirectional ring topology, whose regularity features help in the

¹The lightpaths tear-down can be done independently by the optical layer or upon request from an IP entity, but this does not affect either performances or the overall architecture of the network.

interpretation of results. Besides, rings are among the topologies of choice for the realization of metropolitan area optical networks.

Modified NSFNet — or NSFNet for short (see Fig. 6). This is a modification of the NSF network topology, where we have limited the number of G-OXC's to study whether an irregular, wide area, mesh topology with additional optical resources (the pure OXC's), with respect to the traffic generation points (G-OXC's) does influence the performance of dynamic grooming in overlay IPO.

Table I summarizes the main simulation parameters that are kept constant throughout the simulations unless otherwise stated. Other parameters are defined and explained when needed.

A. 8-Ring topology

One of the wavelengths is reserved to realize a pre-established logical ring topology to ensure connectivity in the data-layer. Some results on this topology were included in [26]; here we focus on the impact of the traffic model and on fairness issues.

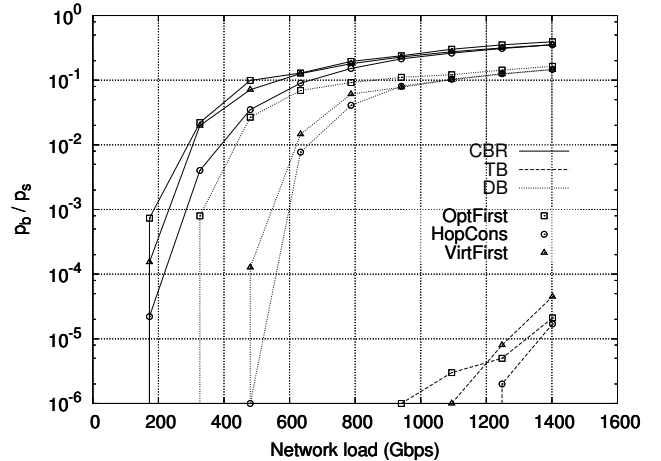


Fig. 2. Blocking probability p_b for the CBR and starvation probability p_s for both the elastic traffic models

Let's first of all analyze the importance of using a realistic traffic model. Fig. 2 presents a comparison between the blocking probability p_b obtained using a circuit-based CBR traffic model (solid lines) with the starvation probability p_s obtained modeling best-effort traffic relations using the TB approach (dashed lines) and the DB approach (dotted lines) when the three grooming algorithms *VirtFirst* (triangle marks), *OptFirst* (square marks) and *HopCons* (round marks) are used. For both elastic traffic models, the minimum requested rate b_m is fixed to 1 Gbit/s.

The difference in performance results yielded by the three approaches is dramatic, showing that the performance

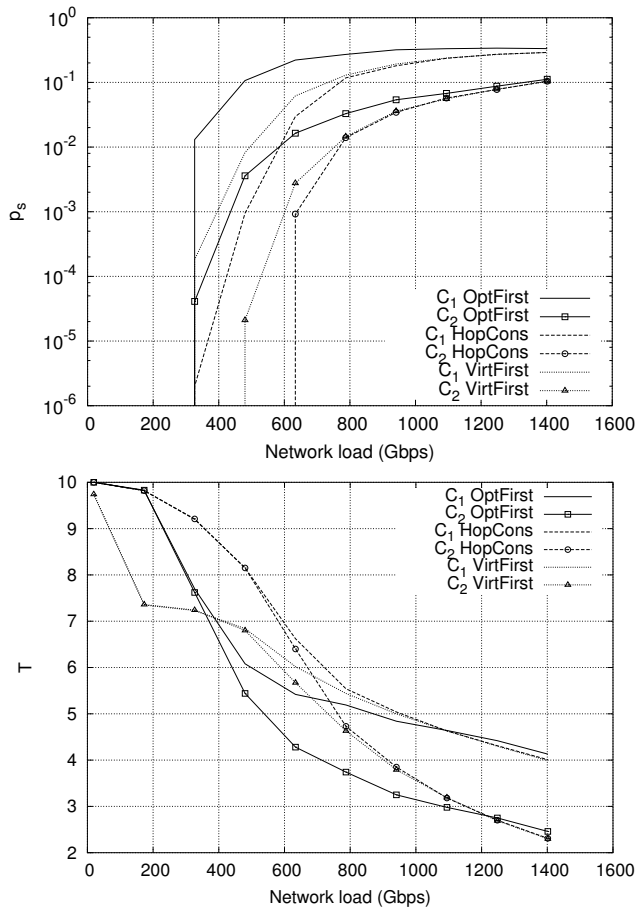


Fig. 3. Evaluating bandwidth requests fairness. Starvation probability p_s (upper plot) and average throughput T (lower plot) users requesting different minimum bandwidth

obtained by the considered grooming algorithms with CBR traffic is very conservative, leading to intolerable levels of blocking probability even for medium to low traffic loads. Instead, when considering a simplistic model of elasticity such as TB, the starvation probability is very low, almost negligible, and this is due to the approximation that a flow request is closed without considering if the data has been completely transmitted or not, thus reducing the actual network load. Therefore, we argue that using a more sophisticated, *data-based* traffic model exhibiting a realistic behavior is essential to evaluate the performance of any dynamic grooming algorithm. From now on we use only the DB traffic model.

We do not include throughput plots because it is useless to compare the throughput of CBR connections (constant!), with the one of elastic traffic. Moreover even the throughput of TB and DB model are not comparable, since the overall amount of transferred data is different.

We now concentrate on fairness issues, analyzing how

different grooming algorithms in overlay IPO networks influence fairness. In particular we investigate two main aspects that influence fairness: the minimum bandwidth b_m required by incoming flow requests, and the physical distance (in terms of the number of crossed OXCs) between the end-nodes of a flow. This latter problem was addressed in [27] in the context of peer IPO networks with a bandwidth-guaranteed traffic model. Resource sharing in general exacerbate the unfairness, because longer flows share the resources with more flows with respect to short ones, thus they remain longer within the network and the more they remain the more are the flows they compete with.

A first set of results regarding the fairness of different grooming algorithms for clients with different SLA requirements is shown in Fig. 3. These results have been obtained differentiating the users in terms of the minimum bandwidth b_m they require to the network to avoid starvation. In particular, we consider two classes of users. Class C_1 with minimum bandwidth request $b_{m1} = 2$ Gbit/s and class C_2 with $b_{m2} = 1$ Gbit/s.

The upper plot shows the starvation probability and the lower one the average throughput. It is clear that the relative merits of grooming policies are unchanged by the presence of classes, but the main result to highlight here is that when considering a realistic traffic scenario, none of the grooming algorithms allows to improve the fairness with respect to the minimum requested bandwidth. The explanation is trivial, but the solution seems to be far more complex, like introducing some form of proportional scheduling in nodes or some form of CAC. In fact, when the traffic shares resources with a max-min fairness criterion, the available network resources are fully shared among all the accepted flows, and it is not possible to distinguish between traffic flows with different minimum bandwidth requirements. In other words, a flow with higher minimum requested rate gets starved with higher probability.

As easily predictable and analyzed in [27], dynamic grooming is prone to an unfair behavior toward user pairs with a longer physical distance in terms of crossed OXCs. A ring is the ideal topology to analyze this behavior, since the regularity of the topology does not introduce any distortion effect. In general the dominating effect is that user pairs distant one another have less chances to setup a new lightpath. An intelligent grooming policy should compensate for, at least in part, this inherent unfairness by *sparing* optical resources to dedicate to longer flows. However, in an overlay model and without costly traffic measurements, this might not be easy to implement. A possible, partial solution, may reside in a more conservative closing procedure for long lightpaths. We have defined a *Length Dependent optical closing (LEDE)* policy that defines the delay of closure for a lightpath of H physical hops as $(H - 1)\tau_{cl}$.

We now study the fairness of the three grooming al-

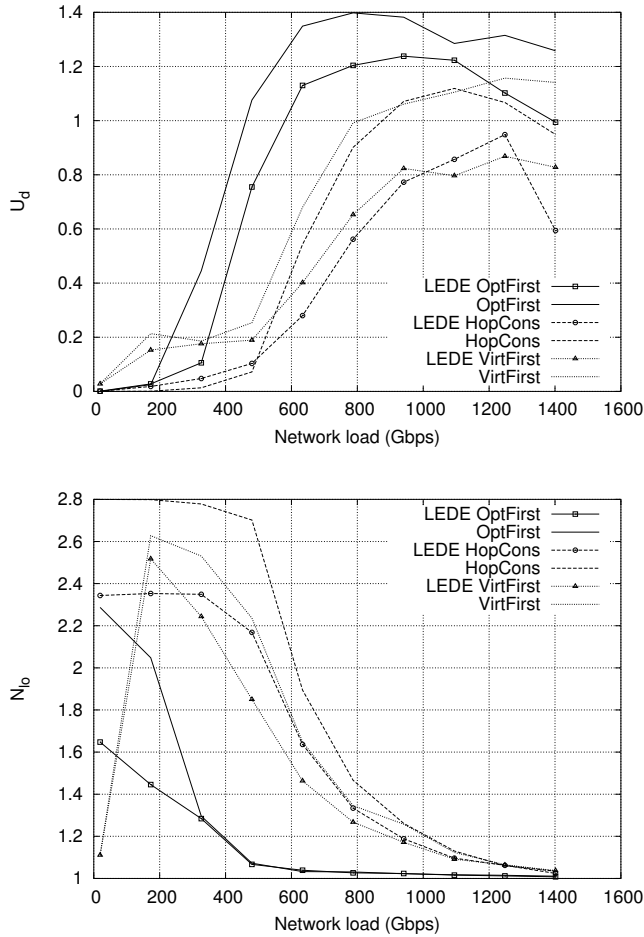


Fig. 4. Comparison for grooming algorithms with and without *LEDE* option: impact over the distance unfairness index U_d (top plot), and average number of links per optical path N_{lo} (bottom plot).

gorithms according to the distance between the end-nodes of a flow, with and without the *LEDE* option as defined above. The base closing timeout τ_{cl} is computed according to the average flow interarrival time as explained in Sect. II-E and it automatically depends on the network load. In each simulation we set $\tau_{cl} = t_{ia}/2$, where t_{ia} is the mean value of the time spent between the arrival of two flow-requests with the same USERS pair as source and destination.

The upper plot in Fig. 4 shows that the *LEDE* option effectively alleviates the unfairness, by uniformly reducing the fairness index. In particular, at very high load (for $p_s \simeq 20\%$), U_d is always below 1 for both *VirtFirst* and *HopCons*.

The impact of the *LEDE* option over the average number of links per optical path is well illustrated in the lower plot, which shows an average increase of N_{lo} for all the grooming algorithms. This behavior proves that increasing the closure timeout on longer optical routes increases the

amount of resources dedicated to longer routes and reduces the unfairness toward longer flows in IPO networks.

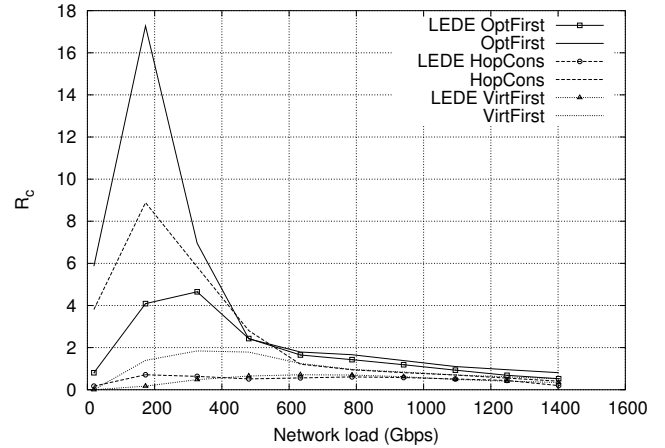


Fig. 5. Comparison for grooming algorithms with and without *LEDE* option: impact over the routing table change rate R_c .

Fig. 5 shows instead a useful ‘side-effect’ of *LEDE*. In fact, keeping lightpaths not carrying active traffic in the network open for longer periods, the frequency of routing table updates due to virtual topology changes is reduced, which is an important cost factor in an IPO network [15]. Starvation probability and throughput (not shown for the sake of brevity) show that this is obtained without average performance losses.

B. NSFNet topology

We consider the NSFNet topology with a mix of OXCs and G-OXCs as shown in Fig. 6. We concentrate on the impact of IP flow requests granularity and the threshold τ_o on the grooming algorithms performance, referring the reader to [13] for additional results on this topology.

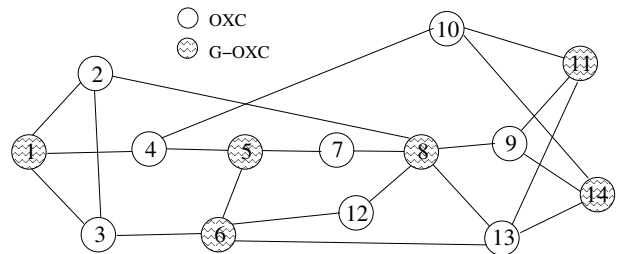


Fig. 6. Topology NSFNet

We first investigate the performance of *VirtFirst* and *OptFirst* as a function of the data-layer requests granularity. Fig. 7 presents T , p_s and p_b of data-layer flows for $B_M = 20, 10$, and 4 Gbit/s, i.e., when requests have a granularity of 1, 1/2, and 1/5, of the wavelength capacity.

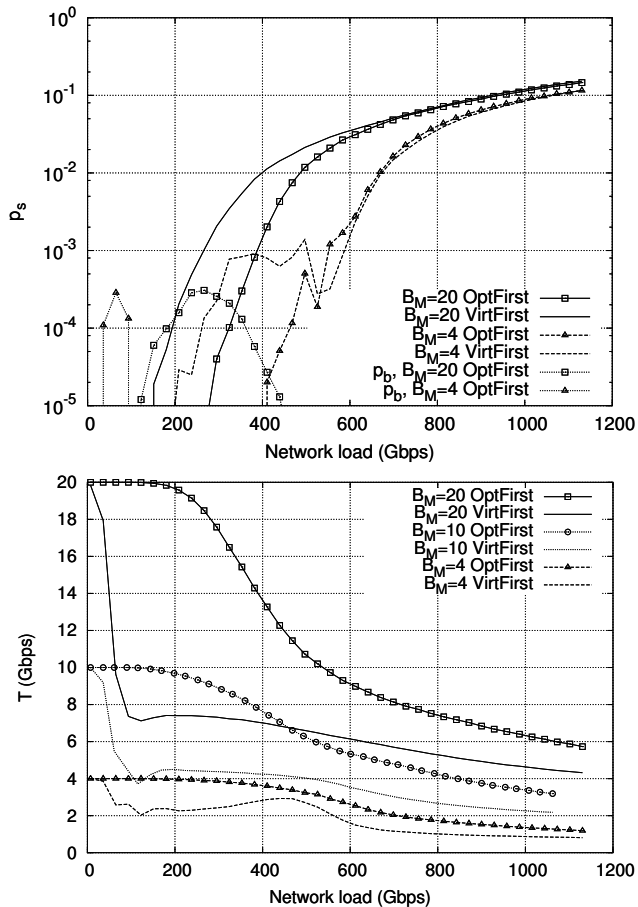


Fig. 7. Throughput T and starvation/blocking probability p_s/p_b as a function of the network load and for different flow granularity

The lower plot shows the throughput T and the upper plot shows the starvation p_s and the blocking p_b . The upper plot reports the curves only for $B_M = 20$ and 4 Gbit/s to avoid cluttering the graph. It is interesting that the absence of a pre-established logical topology does not allow *OptFirst* to ensure $p_b = 0$ at low loads. On the other hand *OptFirst* ensures a higher throughput, but its advantage tends to disappear as the flows granularity decreases. As we did in the 8-Ring topology the “network disconnection” phenomenon can be avoided by pre-establishing a fully connected topology of lightpaths. Notice, however, that in this general mesh topology the choice of the pre-established topology can be non-obvious. For instance a minimum spanning tree can create artificial bottlenecks, while the presence of pure OXC makes other choices, like for instance rebuilding the physical topology at the logical level, far less obvious than in a regular topology like the ring.

The last set of results is related to the impact of the threshold τ_o . Fig. 8 presents the throughput and the starvation

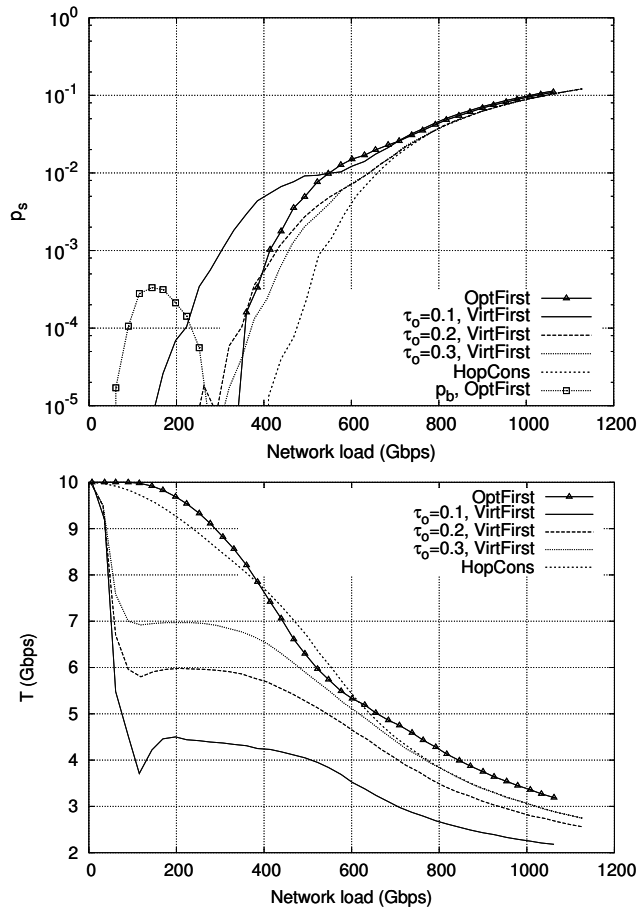


Fig. 8. Throughput T and starvation - blocking probability $p_s - p_b$ as a function of the network load and the threshold τ_o for *VirtFirst* grooming

probability obtained with the *VirtFirst* grooming for $\tau_o = 0.1, 0.2, 0.3$. *VirtFirst* is most sensitive to this threshold, but we also report the curves relative to *OptFirst* and *HopCons* with $\tau_o = 0.3$. *OptFirst* is obviously insensitive to τ_o . Increasing τ_o improves *VirtFirst* throughput and starvation probability as expected, but the performance remains well below *OptFirst* and *HopCons*. Indeed, the effect of limiting the data-layer path with *HopCons* policy seems to be the dominant effect, confirming that this grooming policy has a user perceived quality similar to that provided by *OptFirst*, while limiting the cost for the network.

Fig. 9 reports the average number of hops N_h (upper plot) and the average number of links per lightpath N_{l_o} (lower plot) in the same scenario. The impact of τ_o on these parameters is much smaller than on the average throughput. Finally note that *HopCons* performs much better than *VirtFirst* not only in term of throughput but also when considering the number of links per lightpath, at least for low/medium network loads, as shown in the upper plot of

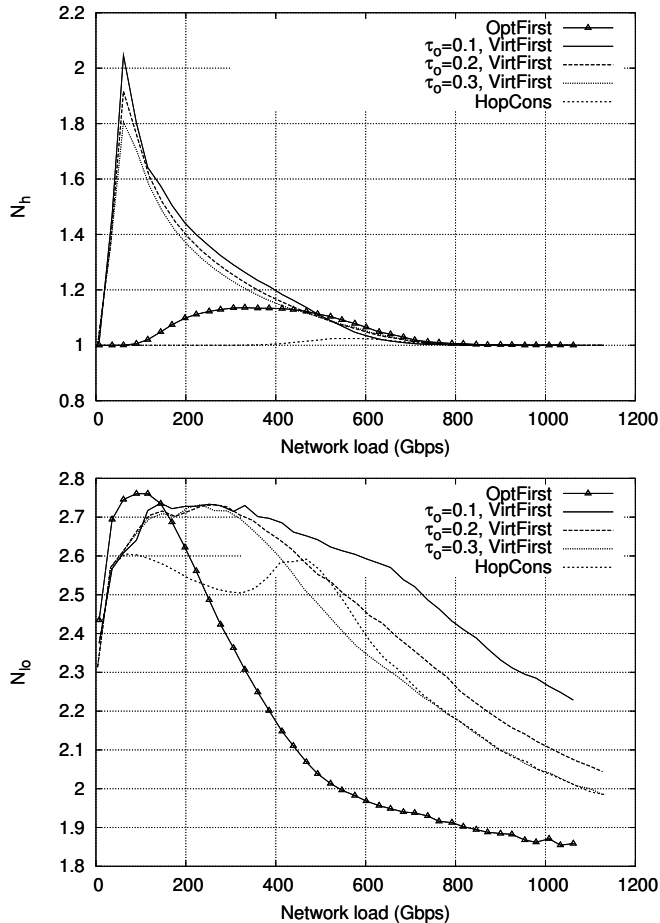


Fig. 9. Average number of hops N_h and average number of links per lightpath N_l for different values of τ_o

Fig. 9. In some cases this grooming policy performs even better than *OptFirst*. The range of the average number of links per lightpath is very compressed as shown by the scale of the lower plot in Fig. 9. The detailed behavior is not easily explained, but the *OptFirst* shows a tendency to open shorter lightpaths at high loads.

V. CONCLUSION & FUTURE WORK

This paper has presented a comparative study of dynamic grooming algorithms in realistic scenarios with a data-based traffic model including elasticity, together with the simulation tool (GANCLÉS) used to perform it.

The first part of the paper is devoted to GANCLÉS presentation, highlighting its innovative features and the management of different architectural models of IPO networks through the explicit simulation of the optical and IP network levels.

The second part of the paper discusses performance results of different dynamic grooming algorithms on two topologies: a ring and a modification of NSFNet. First of

all it is shown how the traffic model impacts on results. Then several performance indices, including the throughput of elastic flows, the probability that the service they receive falls below an acceptable threshold causing the flow starvation, and fairness are compared for the chosen grooming policies assuming an overlay IPO model.

The results show that the use of simplistic traffic model in the design of grooming algorithms may bring to misleading conclusions. Moreover, it is shown that the presence of a double network layer (optical and IP) does not alleviate traditional fairness problems associated with best-effort, elastic traffic. Some form of compensations are possible through the use of smart grooming policies; however, in an overlay model, where no information is shared between the optical and the IP level, it is not easy to find the appropriate and definitive solution.

Further studies on dynamic grooming enabled by GANCLÉS include comparison between different IPO architectures, studying what is the amount of information that needs to be exchanged to allow “intelligent” resource use. In addition, grooming strategies, policies, and algorithms can be implemented and studied in the simulator as we did for the *HopCons* policy that, although very simple, allows overcoming some of the shortcomings of *OptFirst* and *VirtFirst*. Finally, one major question is related to the use of QoS routing either in the optical or IP network layer in order to understand how “intelligent” routing strategies do interact one another through grooming policies.

REFERENCES

- [1] K.H. Liu, *IP Over WDM*, John Wiley, Nov. 2002
- [2] A. Banerjee, J. Drake, J.P. Lang, B. Turner, K. Kompella, Y. Rekhter, *Generalized Multiprotocol Label Switching: an Overview of Routing and Management Enhancements*, IEEE Communications Magazine, 39(1):144–150, Jan. 2001.
- [3] S. Tomic, B. Statovci-Halimi, A. Halimi, W. Mueller, J. Fruehwirth, *ASON and GMPLS: Overview and Comparison*, Photonic Network Communications, 7(2):111–130, March 2004.
- [4] B. Wen, N.M. Bhide, R.K. Shenai, K.M. Sivalingam, *Optical Wavelength Division Multiplexing (WDM) Network Simulator (OWNs): Architecture and Performance Studies*, SPIE Optical Networks Magazine Special Issue on “Simulation, CAD, and Measurement of Optical Networks”, Sep/Oct. 2001.
- [5] H. Zhu, H. Zang, K. Zhu, B. Mukherjee, *A novel generic graph model for traffic grooming in heterogeneous WDM mesh networks*, IEEE/ACM Transactions on Networking, 11(2):285–299, Apr. 2003.
- [6] M. Kodialam, T.V. Lakshman, *Integrated Dynamic IP and Wavelength Routing in IP over WDM Networks*, In Proc. of INFOCOM 2001, pp. 358–366, Anchorage, AK, USA, Apr. 22–26 2001.
- [7] X. Niu, W.D. Zhong, G. Shen, T.H. Cheng, *Connection Establishment of Label Switched Paths in IP/MPLS over Optical Networks*, Photonic Network Communications, 6:33–41, July 2003.
- [8] R. Srinivasan, A.K. Somani, *Dynamic Routing in WDM Grooming Networks*, Photonic Network Communications, 5:123–135, Mar. 2003.

- [9] S. Koo, G. Sahin, S. Subramaniam, *Dynamic LSP Provisioning in Overlay, Augmented, and Peer Architectures for IP/MPLS over WDM Networks*, In Proc. of INFOCOM 2004, Hong Kong, China, Mar. 7–11 2004.
- [10] R. Dutta, G.N. Rouskas, *Traffic grooming in WDM networks: past and future*, IEEE Network Magazine, 16(6):46–56, Nov./Dec. 2002.
- [11] X. Zhang, C. Qiao, *An Effective and Comprehensive Approach to Traffic Grooming and Wavelength Assignment in SONET/WDM Rings*, IEEE/ACM Transactions on Networking, 8(5):608–617, Oct. 2000.
- [12] K. Zhu, B. Mukherjee, *Traffic grooming in an optical WDM mesh network*, IEEE Journal on Selected Areas in Communications, 20(1):122–133, Jan. 2003.
- [13] R. Lo Cigno, E. Salvadori, Z. Zsóka, *Elastic Traffic Effects on WDM Dynamic Grooming Algorithms*, In Proc. IEEE Globecom 2004, Dallas, Texas, USA, Nov. 27 – Dec. 03, 2004.
- [14] D. Bertsekas, R. Gallager, *Data Networks*, Prentice-Hall, 1987.
- [15] B. Rajagopalan, J. Luciani, D. Awduche, *IP over Optical Networks: A Framework* IETF RFC 3717, Mar. 2004.
- [16] GANCLES - Grooming capable Network Call-Level Simulator. URL: <http://netmob.unitn.it/gancles.html>
- [17] ANCLES - A Network Call-Level Simulator. URL: <http://www.tlc-networks.polito.it/ancles>
- [18] C. Casetti, R. Lo Cigno, M. Mellia, M. Munafò, Z. Zsóka, *A Realistic Model to Evaluate Routing Algorithms in the Internet* In Proc. IEEE Globecom 2001, San Antonio, Texas, USA, pp. 1882–85, Nov. 25–29, 2001.
- [19] ASONCLES - ASON Network Call-Level Simulator. URL: <http://www.hit.bme.hu/~zsoka/asoncles>
- [20] K. Pawlikowski, *Steady-state simulation of queueing processes: a survey of problems and solutions*, ACM Computer Surveys, 22(2):123–170, June 1990.
- [21] C. Cavazzoni, V. Barosco, A. D’Alessandro, A. Manzalini et al. *The IP/MPLS over ASON/GMPLS test bed of the IST project LION*, IEEE/OSA Journal of Lightwave Technology, 21(11):2791–2803, Nov. 2003.
- [22] M. Fawaz, B. Daheb, O. Audouin, M. Du-Pond, G. Pujolle, *Service level agreement and provisioning in optical networks*, IEEE Communications Magazine, 42(1):36–43, Jan. 2004.
- [23] S. Sibal, A. DeSimone, *Controlling Alternate Routing in General-Mesh Packet Flow Networks*, in Proc. ACM SIGCOMM’94, London UK, August 1994.
- [24] Q. Ma, P. Steenkiste, and H. Zhang, *Routing High-Bandwidth Traffic in Max-Min Fair Share Networks*, in Proc. of the ACM SIGCOMM’96, pp. 206–217, Stanford, CA, USA, Aug. 1996.
- [25] Z. Wang, J. Crowcroft, *QoS Routing for Supporting multimedia applications*, IEEE JSAC, 14(7):1228–1234, Sept. 1996.
- [26] E. Salvadori, Z. Zsóka, R. Lo Cigno, R. Battiti, *Dynamic Grooming in IP over Optical Networks based on the Overlay Architecture*, Submitted for publication, preprint available: <http://dit.unitn.it/locigno/preprints/SaZsoLoBa04-3.pdf>.
- [27] T. Cinkler, C. Gáspár, *Fairness Issues of Routing with Grooming and Shared Protection*, In Proc. of IFIP ONDM 2004, Gent, Belgium, Feb. 2004.